



Nucleotide

PubMed	Nucleotide	Protein	Genome	Structure	PopSet	Taxonomy	OMIM
Search		Nucleotide	for				
		Limits	Index	History	Clipboard		
Display	Default View	as	HTML	Save	Add to Clipboard		

☐ 1: AL034446 **Streptomyces coelicolor cosmid 1A9** PubMed, Protein, Related Sequences, Taxonomy

LOCUS SC1A9 29400 bp DNA BCT 11-DEC-1998
 DEFINITION Streptomyces coelicolor cosmid 1A9.
 ACCESSION AL034446
 VERSION AL034446.1 GI:4007685
 KEYWORDS ADA-like regulatory protein; araC family; gntR; helix-turn-helix;
 His rich; integral membrane; luxR; luxR family;
 methylated-DNA-protein-cysteine methyl transferase; nucleotide
 binding; ogt; oxidoreductase; poxB; pyruvate oxidase; response
 regulator; sensor kinase; sorbitol oxidase; thiamine
 pyrophosphate; TPP; transferase; transmembrane protein; two
 component system.

SOURCE Streptomyces coelicolor A3(2).
 ORGANISM Streptomyces coelicolor A3(2)
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE 1 (bases 1 to 29400)
 AUTHORS Saunders, D.C. and Harris, D.
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 29400)
 AUTHORS Bentley, S.D., Parkhill, J., Barrell, B.G. and Rajandream, M.A.
 TITLE Direct Submission
 JOURNAL Submitted (04-DEC-1998) Streptomyces coelicolor sequencing project,
 Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
 CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
 David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
 Colney, Norwich, Norfolk NR4 7UH, UK

REFERENCE 3 (bases 1 to 29400)
 AUTHORS Redenbach, M., Kieser, H.M., Denapaite, D., Eichner, A., Cullum, J.,
 Kinashi, H. and Hopwood, D.A.
 TITLE A set of ordered cosmids and a detailed genetic and physical map
 for the 8 Mb Streptomyces coelicolor A3(2) chromosome
 JOURNAL Mol. Microbiol. 21 (1), 77-96 (1996)
 MEDLINE 97000351

COMMENT Notes:
 Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL; http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).
 The more significant matches with motifs in the PROSITE database
 are also included but some of these may be fortuitous. The length
 in codons is given for each CDS.
 Usually the highest scoring match found by fasta -o is given for
 CDS which show significant similarity to other CDS in the database.
 The position of possible ribosome binding site sequences are given
 where these have been used to deduce the initiation codon. Gene
 prediction is based on positional base preference in codons using a
 specially developed Hidden Markov Model (Krogh et al., Nucleic

BEST AVAILABLE COPY

Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.go.jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 1A9 lies between 9B2 and 6C5 in the Ase-I-B genomic restriction fragment.

FEATURES	Location/Qualifiers
source	1..29400 /organism="Streptomyces coelicolor A3(2)" /strain="A3(2)" /db_xref="taxon:100226" /clone="cosmid 1A9"
<u>gene</u>	complement(1..651) /gene="SC1A9.01c"
<u>CDS</u>	complement(<1..651) /gene="SC1A9.01c" /note="SC1A9.01c, incomplete CDS, possible transferase, partial CDS, len: 219 aa, similar to TR:O53185 (EMBL:AL021246) a proposed transferase from Mycobacterium tuberculosis (302 aa), fasta scores: opt: 764, z-score: 1097.3, E(): 0, (62.1% identity in 203 aa overlap (302 aa). Also weakly similar to several methyltransferases eg. METH ECOLI (EMBL:X16584) 5-methyltetrahydrofolate-homocysteine methyltransferase (1226 aa), fasta scores; opt:160, z-score: 234.8, E(): 8.4e-06, (30.0% identity in 220 aa overlap)." /codon_start=1 /transl_table=11 /product="putative transferase" /protein_id="CAA22371.1" /db_xref="GI:4007686" /translation="MTSDFADALASGPLVLDGGLSNQLEAAGHDLDGALWSARLLAED PEAITRAHLAYFEAGAEVAITSSYQATFEGFARRGIGRERAAELLALSVASAREAAARR ARTARPERALWVAASAGPYGAMLADGSEYRGYGLGRGALERFHRPRLEVLAAARPDV LALETVPDTEAAAALLRAVRGLDVPWLSYTVAGDRTRAGQPLDEAFALAADVDEVI"
<u>gene</u>	731..1654 /gene="SC1A9.02"
<u>CDS</u>	731..1654 /gene="SC1A9.02" /note="SC1A9.02, possible transmembrane protein, len: 307aa; Contains several possible membrane spanning domains." /codon_start=1 /transl_table=11 /product="putative transmembrane protein" /protein_id="CAA22372.1" /db_xref="GI:4007687" /translation="MTISGRIRSVRCSPRHSSHGPDGVPVRFVWQFLAVLVAYAIG GIAVQAVKDNWLTLLVVGLTVALVVFVYAWVVRTERREALDVALDGAAGAKAGWGTLL IGFGLFGAVITNLFASGYEVDGLGSGVQGAIGLVGFMAAAAATEEVVFRGVLFRIIEE HIGTYLALGLTGLVFGLMHLLNEDATLWGALAIIEAGFMLAAAYAATRNLTIGVH FGWNFAAGGVFSTVVSNGDSEGLLDATMSGPKLLTGDFGPEGSVYSVGFVLLTLV FLWLAHRRGNIVAFGSRRRAAGANSAATLPR"
<u>gene</u>	1651..2817 /gene="SC1A9.03"
<u>CDS</u>	1651..2817

/gene="SC1A9.03"
/note="SC1A9.03, probable two component sensor kinase,
len: 388aa; similar to a family of sensor kinases egs.
TR:Q53893 (EMBL:U51332) AbsA1 from Streptomyces coelicolor
(571 aa), fasta scores; opt: 389, z-score: 278.9, E():
2.9e-08, (33.1% identity in 405 aa overlap) and UHPB_ECOLI
(EMBL:M17102) sensor kinase from Escherichia coli (500
aa), fasta scores; opt:221, z-score: 248.0, E(): 1.5e-06,
(25.9% identity in 398 aa overlap). Contains several
possible membrane spanning domains."
/codon_start=1
/transl_table=11
/product="putative sensor kinase"
/protein_id="CAA22373.1"
/db_xref="GI:4007688"
/translation="MIDRRRVLELWRRDLVTVRDLPGLVLLLVASLLPSLRGQGTEIG
GLPTRPADALAGVAAVLQSIPLAVRRRWTLCLTLVSLGFALDQLRAYHLFAGAALPI
VLINAGSHQEKYRRATQVTATLGYVAMAVGLNARGGDETLVEYVTFYLVLALAWGIGA
WMRSARAAEAERRSRVAEDARNAERTRIARELHDVVTHVVTAMVVQSEAARYLTAAPE
RLDESLAAVSDTGRRAITDLRHLLDLLNPDHGTAEPRTPPVGRVLTLEQTRRAGQPV
EFTEEGTPAAATGSSDLVAYRVVQEALTNALKYDHGGRTSVLVRHGEREITVEVGTDG
SGSGAASPGGSGRGLAGLRERVDVLGGFSTDRPADGGFVVRARIPGGSGGSTA"
2814..3473
gene
/gene="SC1A9.04"
2814..3473
CDS
/gene="SC1A9.04"
/note="SC1A9.04, probable luxR family response regulator,
len: 219 aa; similar to a family of regulators egs.
TR:O69816 (EMBL:AL023496) probable two-component regulator
from Streptomyces coelicolor (224 aa) fasta scores;
opt:714, z-score:1103.5, E():0, (54.8% identity in 219 aa
overlap) and NARL_ECOLI (EMBL:X13360) nitrate /nitrite
response regulator from Escherichia coli (216 aa) fasta
scores; opt: 465, z-score: 656.5, E(): 2.7e-29, (37.6%
identity in 213 aa overlap). Contains Pfam match to entry
PF00072 response_reg, Response regulator receiver domain,
score 111.10, E-value 2.1e-29, Pfam match to entry PF00196
GerE, Bacterial regulatory proteins, luxR family, score
79.50, E-value 7e-20, PS00622 Bacterial regulatory
proteins, luxR family signature and an helix-turn-helix
motif from: 1 to: 219, Score 983 (+2.53 SD)."
/codon_start=1
/transl_table=11
/product="putative response regulator"
/protein_id="CAA22374.1"
/db_xref="GI:4007689"
/translation="MSAPIRVVICDDQALIRTGLATIVDAQPDLEVVGECGDGQTGVD
LARELRPDVVVMDIRMPVLDGLEATRLLAGAGVAHPVKVLVVTTFNLDYVYEALRAG
ASGFLLKDAPPDRLLHGIRTVAMGAALLDPDVTRRLVGRYAARIRPAEGTARDIPLTP
RETEVLRILIADGLSNSEIAAALVISPETVKTFVSRILTKLDLRDRVQAVVFAYRHGLV
T"
misc feature
2826..3173
/gene="SC1A9.04"
/note="Pfam match to entry PF00072 response_reg, Response
regulator receiver domain, score 111.10, E-value 2.1e-29"
misc feature
3276..3470
/gene="SC1A9.04"
/note="Pfam match to entry PF00196 GerE, Bacterial
regulatory proteins, luxR family, score 79.50, E-value
7e-20"
misc feature
3327..3410
/gene="SC1A9.04"
/note="PS00622 Bacterial regulatory proteins, luxR family
signature."
gene
complement(3492..4094)

CDS

```
/gene="SC1A9.05c"  
complement(3492..4094)  
/gene="SC1A9.05c"  
/note="SC1A9.05c, possible transmembrane protein, len: 200  
aa; contains possible membrane spanning hydrophobic  
regions"  
/codon_start=1  
/transl_table=11  
/product="putative transmembrane protein"  
/protein_id="CAA22375.1"  
/db_xref="GI:4007690"  
/translation="MFRANAGIRSAPYVHVFPMRGDMASSTDRPRLRRRDSLWGI  
GVMALLAGVVVRLVLNGTSAWLSALLGAVPAAVWVWVRRRRVRDARAVGAEPDDVP  
AMERQILKGGPAPRDPERRRAMAAAFVDSRQERLRRNRWAFPMPLAVIFFGTSALWYLS  
GSVGAGSLMLGLGVVFLGWLAWYNLRIDRRLSHMRGRLRG"
```

gene

4140..4631

CDS

```
/gene="SC1A9.06"  
4140..4631  
/gene="SC1A9.06"  
/note="SC1A9.06, unknown, len: 163 aa"  
/codon_start=1  
/transl_table=11  
/product="hypothetical protein SC1A9.06"  
/protein_id="CAA22376.1"  
/db_xref="GI:4007691"  
/translation="MRDMNDTRPTSTVTSTRTPADAVTGMVDHVLHLAAGWTRWDGT  
PAHVDGRVYTPHKAIRRVADHLVDHLAELEARLAGEETQPDHWHASLVTTEADRAAFT  
AEDLDEARSRLTRLARIWANRLDALTDQLDHSPGEGWSFRELAHHAESGYADAVG  
DLS"
```

gene

4628..5383

CDS

```
/gene="SC1A9.07"  
4628..5383  
/gene="SC1A9.07"  
/note="SC1A9.07, unknown, len: 251 aa; similar to  
TR:O33986 (EMBL:U82823) hypothetical protein from  
Saccharopolyspora erythraea (266 aa), fasta scores; opt:  
740, z-score: 869.0, E(): 0, (53.4% identity in 251 aa  
overlap)."  
/codon_start=1  
/transl_table=11  
/product="hypothetical protein SC1A9.07"  
/protein_id="CAA22377.1"  
/db_xref="GI:4007692"  
/translation="MTATAFAALHRAGEPLLLPCAWDHASAFALAGQGFRVGTTSLG  
VAAAAGLPDGASATRDETLRLALVLGSAPFPLSVDAEDGFSDDPDEVGEFARQLAAVG  
AVGINLEDGLGPVGRHAAKIAAVRSAAPGLFVNARTDTYWSGDGDVTETLRRLEAYRE  
AGADGVFVPGLTDPARIGSLAARFDVPLNVLYTPAGPGLAHLADLGVRRVSLGSLLYR  
RALGAALRAAADVRAGRDPGGPTPTYDEVRAPG"
```

gene

complement(5386..5955)

CDS

```
/gene="SC1A9.08c"  
complement(5386..5955)  
/gene="SC1A9.08c"  
/note="SC1A9.08c, possible transcriptional regulator, len:  
189 aa; similar to TR:O34892 (EMBL:AF027868) proposed  
transcriptional regulator from Bacillus subtilis (191 aa)  
fasta scores; opt: 302, z-score: 488.2, E(): 6.6e-20,  
(31.8% identity in 173 aa overlap). Contains an  
helix-turn-helix motif from: 1 to: 189, Score 1074 (+2.84  
SD)."  
/codon_start=1  
/transl_table=11  
/product="putative transcriptional regulator"  
/protein_id="CAA22378.1"  
/db_xref="GI:4007693"  
/translation="MPRVGLTTDRVVAAAADLADETGFSVTVSALARHFGVKDASLY"
```

THVRNLQDLRVRVALLAGGELIEE IAGAVAGRAGKEALAAFAGAYRAYALRHPGRYAA
TQIRVDQSLVADSAALRRTAETIYGM LRSYGLTEPDLTDAVRLLRSTFHGYCALESSG
AFGAPRDVRASWDKAVDALHVALENWPRA"

gene 6107..7825
/gene="SC1A9.09"

CDS 6107..7825
/gene="SC1A9.09"
/note="SC1A9.09, unknown, len: 572 aa; Contains His rich
extreme N-terminal."
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1A9.09"
/protein_id="CAA22379.1"
/db_xref="GI:4007694"
/translation="MGHGHGHPHGHHHHGHGTHDHEHATEQVLPAAFDTSVPDEALSPA
QQSRRGLLRAGLLGAGLAAGTVLAPAATATAAPARAASNGRRGKGF L WLAGDHHIHT
QYSSDGKYRVVDQVRQGARHGM DLVITDHGSNTHAKIGVEKVNPDIREARAAHEDTL
VFQGLEWNI PA AEHGT V FVHPGKHEVSVLKQFETDYDGSVKAGDSTPANEALAIAGL
SFLADQVKRRKVKDALMLANHPARKGIDSPHEIRAWRDATSRGHQIAVGFEGAPGHQA
GGLPEPLGPGGARGIYDGSPSANSFAGYPLESYRTWGGFDWMTATVGGLWDSLIAEGR
PWWITANSDSHQVYADTGARGGGDFNANGRYDDPVYAGQIDITQNDFWPGQYSRTHVG
SDGFSYAAVMDGIRAGRVVWDHGLVSGLDVRVSGGGRWATLGGALHVRGTRVTL SI
DVALAGGPNWAGFV PKLARVDVIQGDVTGPAADKDTFTAPTARVVKSYEVDKETGTVR
LTYDLGRVDRPVYLRTRGTDGNRS AVGSLGAKVDPAGPAIDVVG DADPWRDLWFYSNP
VWVLPS"

gene 7822..8346
/gene="SC1A9.10"

CDS 7822..8346
/gene="SC1A9.10"
/note="SC1A9.10, unknown, len: 174 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1A9.10"
/protein_id="CAA22380.1"
/db_xref="GI:4007695"
/translation="MTPPAAPYVLGIDTDATT L READHLLQALAAELDLPEGVFGCTH
LVRDGRPRVALSLAAEAEPVLR TARDRLTARGHEVRDGTWDESGRAVLFPGAAALTGT
LTLAELLARS AVDRVTVLGTPDEPSPDTRLVTRNHVRPHWQDGR L VLAAMPAVG GTLV
PFEDPDPTPCCADH"

gene complement(8377..9633)
/gene="SC1A9.11c"

CDS complement(8377..9633)
/gene="SC1A9.11c"
/note="SC1A9.11c, oxidoreductase, len: 418 aa, similar to
many eg. TR:P97011 (EMBL:AB000519) proposed sorbitol
oxidase from Streptomyces sp. (420 aa) fasta scores; opt:
1544, z-score: 1630.2, E():0, (60.0% identity in 413 aa
overlap) and TR:O50531 (EMBL:AL009204) FAD-dependent
oxidoreductase from Streptomyces coelicolor (445 aa) fasta
scores; opt: 374, z-score: 499.4, E(): 1.6e-20, (31.4%
identity in 439 aa overlap)."
/codon_start=1
/transl_table=11
/product="putative oxidoreductase"
/protein_id="CAA22381.1"
/db_xref="GI:4007696"
/translation="MSDITVTNWAGNITYTAKELLRPHSLDALRALVADSARVRVLGS
GHSFNEIAEPGDGGVLLSLAGLPSVVDVDTAARTVRVGGGVRYAELARVVHARGLALP
NMA SLPHISVAGSVATGTHGSGVNGSLASVVREVELVTADGSTVVIARGDERFGGAV
TSLGALGVVTS L TLDLEPAYEMEQHV FTELPLAGLDPATFETVMAAAYSVS LFTDWRA
PGFRQVWLKRRTDRPLDGF PYAAPAAEKMHVPVPGMPAVNCTEQFGVPGPWHERLPHFR
AEFTPSSGAELQSEY LMPREHALAALHAMDAIRETLAPVLQTCEIRTVAAADAQWLS PA
YGRDTVA AHFTWVEDTAAVL PVVRRLEEALVPFAARPHWGKVFTVPAGELRALYPRLA
DFGALAGALDPAGKFTNAFVRGVLAG"

gene 9740..11068

CDS /gene="SC1A9.12"
9740..11068
/gene="SC1A9.12"
/note="SC1A9.12, unknown, len: 442 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1A9.12"
/protein_id="CAA22382.1"
/db_xref="GI:4007697"
/translation="MDSRTALVEDLMERFPHVPREAVFKEDLLRGGVAFDPSALSDDT
NEAAGEVKPKSYFIFSFHDGTLPELGEAALRRPPEEIIITGGPYDLRRTVSVRVNPA
SPYRVAANEDGVLGLYLDGKRIADVGVPPMPEYYRHKLSNGKSVMEVAPTIQWGYLIY
LTAFRVCQYFGAKEECQYCDINHNWRQHKAAGRPTYGVKDVDEVLEALEIIDKYDTAK
ISTAYTLTGGAITSKVQGLDEADFYGRYAKAIEEHFPGRWIGKVVAQALPKPDVQRFK
DYGVDQIYHPNFEVWDEYLFKMYCPGKERYVGRDEWHKRILDSTEVFGARNVIPNFVAG
VEMAEPFGFKTVDEAIESTTEGLRFFFMHSHGITPRFTTWCPEPTTPLGKTNPDGAPLEY
HIRLLQAYRQTMEDYGLSSPPGYGPPGAGNAVFSVSSFMDSLDPEDAPVEV"

gene 11316..12416

CDS /gene="SC1A9.13"
11316..12416
/gene="SC1A9.13"
/note="SC1A9.13, possible nucleotide binding protein, len:
366 aa; similar to several hypotheticals eg. YJEQ_HAEIN
(EMBL:U32844) hypothetical protein HI1714 from Haemophilus
influenzae (346 aa) fasta scores; opt:472, z-score: 564.4,
E(): 3.7e-24, (31.0% identity in 313 aa overlap). Contains
PS00017 ATP /GTP-binding site motif A (P-loop)."
/codon_start=1
/transl_table=11
/product="putative integral membrane nucleotide binding
protein"
/protein_id="CAA22383.1"
/db_xref="GI:4007698"
/translation="MTSTSSHSNHSALSSYGWDDSWADAFAPYAAEGLLPGRVVRVDR
GQCDVVTADGVLRADTAFVTPHDLRVVCTGDWVAVEPGGNPRYVRTYLPRRTAFVRS
TSSKRSEGQILANVDHAVVAVSLAVELDLARIERFLALAWESGAQLVVLTKADLVP
DPVTLAYLVQDVETAAPGVPLPVSAEQEGGLDVLAAVVSOGGTAVLLGQSGAGKSTLA
NALLGEAAMDVQAIRVDVGKGRHTTTTRNLLALPGGGVLIDTPGLRGVGLFDAGNGVD
QVFAEIAELAEECRFHDCAHSESEPGCAVLAAIDSGALPERRLESYRKLRENQRIVAK
TDARARAEIRKEYKRRGAIGKAAMEAKRGGLR"

misc feature 11937..11960

gene /gene="SC1A9.13"
/note="PS00017 ATP /GTP-binding site motif A (P-loop)"

CDS 12489..13922
/gene="SC1A9.14"
12489..13922
/gene="SC1A9.14"
/note="SC1A9.14, possible ADA-like regulatory protein,
len: 477 aa; similar to ADA_MYCTU (EMBL:Z73902) putative
ADA regulatory protein from Mycobacterium tuberculosis
(496 aa) fasta scores; opt: 1135, z-score: 1376.3, E(): 0,
(51.6% identity in 486 aa overlap). Also note N-terminal
region similar to N-terminal region of ADA_ECOLI
(EMBL:M10211) ADA regulatory protein from Escherichia coli
(354 aa) fasta scores; opt: 302, z-score: 255.8, E():
5.7e-07, (34.5% identity in 171 aa overlap) and C-terminal
region similar to 3MG2_ECOLI AlkA, 3-methyladenine DNA
glycosylase II involved in DNA repair (282 aa) fasta
scores; opt: 171, z-score: 232.9, E(): 1.1e-05, (30.0%
identity in 290 aa overlap). Contains Pfam match to entry
PF00165 HTH_2, Bacterial regulatory helix-turn-helix
proteins, araC family, score 72.90, E-value 6.7e-18 and
PS00041 Bacterial regulatory proteins, araC family
signature."
/codon_start=1

/transl_table=11
/product="putative ADA-like regulatory protein"
/protein_id="CAA22384.1"
/db_xref="GI:4007699"
/translation="MTPQTVQPAEHADAREDVRYEAVRSRDARFDGAFFFAVETTGIIY
CRPSCPAVTPKRRNVRRFFATAAAAQSGSFRACRRCPDAVPGSADWNVRADVGRAMR
LIGDGVVDREGVAGLAGRLGYSARQVQRQLTAEVGAGPVALARAQRAHTARVLLQTTV
LPVTEIAFASGFASVRQFNDRIRAVYAATPSELRAAAPARDRAARRTATPSAGVPLRL
AHRGPYQAGPVFDLLQREAVTGVEEVSGETGRRLYRRTLRLPYGTGIVAVQERPGRAG
TGSGGWLEARLHLTDLRLTTSVQRLRRLFDLDADPYAVDERLGADPRLAPLVAARPG
LRSPGTADPAELAVRALVGRTEAERLVQRYGKALDAPCGTLTHLFPEPDVLGAAPHG
TPGALAAALADGAVRLDPGADRDDAERALLAVPGLDARTVAVVRTRALGDPDVAPPGA
AVPDTWRPWRSYALNHLRAAGEWENDR"
misc feature 12819..13079
/gene="SC1A9.14"
/note="Pfam match to entry PF00165 HTH_2, Bacterial
regulatory helix-turn-helix proteins, araC family, score
72.90, E-value 6.7e-18"
misc feature 12930..13058
/gene="SC1A9.14"
/note="PS00041 Bacterial regulatory proteins, araC family
signature."
gene 13919..14479
/gene="SC1A9.15"
CDS 13919..14479
/gene="SC1A9.15"
/note="SC1A9.15, ogt, methylated-DNA-protein-cysteine
methyltransferase len: 186 aa; similar to many eg.
DAT1_BACSU (EMBL:X15659) methylated-DNA-protein-cysteine
methyltransferase from Bacillus subtilis (165 aa) fasta
scores; opt:357, z-score: 478.2, E(): 2.3e-19, (41.4%
identity in 157 aa overlap). Contains Pfam match to entry
PF01035 Methyltrans, 6-O-methylguanine DNA
methyltransferase, score 170.70, E-value 2.5e-47."
/codon_start=1
/transl_table=11
/product="putative methylated-DNA-protein-cysteine
methyltransferase"
/protein_id="CAA22385.1"
/db_xref="GI:4007700"
/translation="MTTTPPTTTTTSIPAETYWHEVDSPVGPLLLTAGSDGALTSLSV
PGQKGGRSVRDGRHDAGPFRVAEEQLGAYFAGELTEFSLPLRAQGTAFRRERVAAALD
DVPYGATTTYGEIAARIGASRPRAVAVGGAIGANPLLLLRPCHRVIGADGSLTGYAGG
LERKTRLLSLEGAPLSRPVPLPATPR"
misc feature 14111..14446
/gene="SC1A9.15"
/note="Pfam match to entry PF01035 Methyltrans,
6-O-methylguanine DNA methyltransferase, score 170.70,
E-value 2.5e-47"
gene complement(14442..14924)
/gene="SC1A9.16c"
CDS complement(14442..14924)
/gene="SC1A9.16c"
/note="SC1A9.16c, small hydrophobic protein, len: 160 aa"
/codon_start=1
/transl_table=11
/product="small hydrophobic protein"
/protein_id="CAA22386.1"
/db_xref="GI:4007701"
/translation="MGAWDLALLAGLVILLGLCGVLLPGVPGSWLVWAGVLWWALKDPR
PLAWAVLVGSTVVLLLSRAVRWALPTRRKRRDEAMRRLTAYAGAGAVLGFLVPLVGA
VPGFMGGIYLAERLRLGRHGEAMASLRTAMRQGGADLLTELFACLLITGAWLGAVLAG
"
gene complement(14986..15969)
/gene="SC1A9.17c"

CDS

```
complement(14986..15969)
/gene="SC1A9.17c"
/note="SC1A9.17c, possible regulatory protein, len: 424
aa; similar to TR:Q56084 (EMBL:D63904) ESA36 protein of
undefined function from Streptomyces thermoviolaceus (334
aa) fasta scores; opt: 313, z-score: 401.9, E(): 4.1e-15,
(30.0% identity in 337 aa overlap). Also similar to
SW:BRPA_STRHY bialaphos biosynthetic pathway regulatory
protein from Streptomyces hygroscopicus (256 aa), fasta
scores; opt: 234, z-score: 233.8, E(): 9.5e-06, (30.1%
identity in 163 aa overlap). Contains helix-turn-helix
motif from: 1 to: 424, Score 1064 (+2.81 SD). Possible
coiled-coil from 171 to 200 (30 residues) Max score: 1.447
(probability 0.87)"
/codon_start=1
/transl_table=11
/product="putative regulatory protein"
/protein_id="CAA22387.1"
/db_xref="GI:4007702"
/translation="MLGVIGLEDTHEAAYRALVSVGAADVDPDLARRLALGERDTERAL
RRLEQNGLAQSSARPGRWVAAPPGVALGALLTQQRHELERAELAAALLAEEYRAAAA
EPAVHDLVEVVTGAGAVAQRFLQLQLGASEEVCALVTDKPVAVTGMENDAEEQATGRG
VRYRVVVVERSVDLPTGITELTAALGRDEQVRVVDVPTKLIVADRSALVPLTARSS
EPAALVVHASGLLELLCGLFEAVWRDALPLRLGASGVTEQAPDGPDGTDLIELSLLLA
GLTDASVAKQLDLGLRTVQRRVKRLMELTGVTTTLQLGWHAYERDWWARRD"
```

gene

16892..18064

CDS

```
/gene="SC1A9.18"
16892..18064
/gene="SC1A9.18"
/note="SC1A9.18, unknown, len: 390aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1A9.18"
/protein_id="CAA22388.1"
/db_xref="GI:4007703"
/translation="MSSPASGAGRPAAAPRRPDPRGPTTAGSGAADLREPPEAVPPPRT
LPDLPLPPRFLPSTEADRDYVGSGPPTYDAEPTALPPADPDGLDGLVPDVTVLEGARY
GTCTLRVSVRGDSARYRGEPDRDALLVARFGAGEQALVLVAMATGARATAGAHRAAA
EVCRWIGRAVGRSHARLAEDLRAARRGDLKSGLHRLTDRSLGRRLRAGAAEQGLAPDEY
AATLRCLLLPADPGCRTRVFFGVGAGGLRLRDGAWQDMEPDAGDVTGEPVLGFGSAP
RDAPRDAPRETPEDDRLLTMDLGITTSPPGYEGPPAGPPREPFRFRASVARPGDVLLMC
TAGLAEPPLLSEPGLELLARRWAARPAPGPGEFLADSGVRVVKGYADDRTA AA VWEA"
```

gene

18093..19835

CDS

```
/gene="SC1A9.19"
18093..19835
/gene="SC1A9.19"
/note="SC1A9.19, poxB, pyruvate oxidase, len: 580aa;
similar to many including POXB_ECOLI (EMBL:X04105) PoxB,
pyruvate oxidase from Escherichia coli (572 aa) fasta
scores; opt:2086, z-score: 2458.4, E():0, (53.1% identity
in 571 aa overlap). Contains Pfam match to entry PF00205
TPP_enzymes, Thiamine pyrophosphate enzymes, score 499.70,
E-value 2.4e-154."
/codon_start=1
/transl_table=11
/product="pyruvate dehydrogenase"
/protein_id="CAA22389.1"
/db_xref="GI:4007704"
/translation="MAKQNVAEQFVDILTRAGVERLYGVVGDSLNPVVDVRRHSGIE
WVHVRHEETA AFAAGAEAQITGKLTACAGSCGPNLHLINGLYDAHRSMAPVLALASQ
IPSEIGLGFFQETHPDQLFRECSHYSELISPKQMPRLQLTAIQHAVGQGGVSVVSL
PGDIADEPAPQGAETALVTSRPTVRPGDEEIDRLVRMIDDADKVTLCGSGTAGAHA
EVMEFAGKLKAPVGHALRGKEFIQYDNPYDVGM SGLLG YGAAYEATHECDLLLLIGTD
FPYNAFLPDDVKIAQIDVRPEHLGRRSKLDLAVWGDARETLRCLIPRVKEKKNRRFLD
RMLKKHADALEGVVKAYTRKVDKHPVPIHPEYVAALLDEMADDDAVFTVDTGMCNVWAA"
```


RYISPNGRRRIIGSFSGHSMANALPMAIGAQFTDRRRQVVSMGDDGGFTMLMGDFLTL
VQHDLPVKIVLFNNSSLGMVELEMLVAGLP SHGVANKNPDAFAVAEACGAFGVRVEKP
KDLAGALKAAFKHKGPALVDVVTDPNALSIPPKISADMVTGFALSASKIVLDGGVGRM
LQMARSNLRNVPRP"

misc feature 18132..19679
/gene="SC1A9.19"
/note="Pfam match to entry PF00205 TPP_enzymes, Thiamine
pyrophosphate enzymes, score 499.70, E-value 2.4e-154"

gene 19979..20386
/gene="SC1A9.20"

CDS 19979..20386
/gene="SC1A9.20"
/note="SC1A9.20, possible regulatory protein, len: 135aa;
similar to TR:O69204 (EMBL:U33059) hypothetical protein
from Amycolatopsis mediterranei S699 (144 aa) fasta
scores; opt:225, z-score:335.1. E(): 2.2e-11, (41.7%
identity in 115 aa overlap). Also similar to TR:Q53897
(EMBL:X60316) AbaA regulatory locus for antibiotic
production in Streptomyces coelicolor (192 aa), fasta
scores; opt: 155, z-score: 284.6, E(): 1.4e-08, (33.3%
identity in 120 aa overlap)"
/codon_start=1
/transl_table=11
/product="putative regulatory protein"
/protein_id="CAA22390.1"
/db_xref="GI:4007705"
/translation="MPQLRRRLGRADLRVPEARRALRELLRHWGGPGQSEVAELLAS
ELVTNALVHTDEGAVLTATVGPRALRVEVRDFVGRGRQPRPRAPQREESTNGRGLVLV
ESLADDWGVQPCEVGKSVWFELGAGAEAEAEAA"

gene complement(20444..21508)
/gene="SC1A9.21c"

CDS complement(20444..21508)
/gene="SC1A9.21c"
/note="SC1A9.21c, possible transmembrane protein, len:
354aa; contains possible membrane spanning hydrophobic
domains in the N-terminal region and possible coiled-coil
from 221 to 250 (30 residues) Max score: 1.449
(probability 0.87)."
/codon_start=1
/transl_table=11
/product="putative transmembrane protein"
/protein_id="CAA22391.1"
/db_xref="GI:4007706"
/translation="MRLTDISLNWLLPGAVLLLGMLAAVAVLARGKRSSGKDAGADDS
WERMEERRRRKEALYGTFSYVLLFCCAATAAALS FHLVGFGEQNLGLSDGWQYLVPF
GLDGAAMFCSVLAVREASHGDAALGSRILVWAFATAA WFNWVHAPRGLGHAGAPHFF
AGMSLSAAVLFDRA LKQTRRAALREQGLVPRPLPQIRMVRWLRAPRETYRAWSLMLLE
GVRSLDEAVEEVRRDRRQKEEKKLRRREQERLER AQLKAISRGHGHRGFPGRGGRQVE
VEVQQVERGSERAAAEPAISTPEQLPAASRRPSLQPVRSGSEQMSVGTVDLTAEDDTQ
ALPRLD SLERKLDLEQQFG"

gene complement(21766..22611)
/gene="SC1A9.22c"

CDS complement(21766..22611)
/gene="SC1A9.22c"
/note="SC1A9.22c, unknown, len: 281aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1A9.22c"
/protein_id="CAA22392.1"
/db_xref="GI:4007707"
/translation="MPVARTALTDAYTRLSEVL PGLGVTELA AAEVPSGGGWVTAAS
LAAGGTELA AFLAWDEAQVLRDYGQRRPDIAS FGLHRYAWPACLLITVPWFLHRRV
PRHPAAHVS YDRTAAGLPLGRMAVRAASFACLP GDPAALPGARVVADEEALRAEVRA
AVAEHLEPVL AGFGPRMRRRGRALWGMATDEVVEGLWYVAHLLGEQERARHELELLLP
GATRPYVGKAAFRELKGP DGAPLHTRDRASCCMFYTLRPEDTCATCPRTC DADRVDKL

gene LAAAG"
22780..23601
/gene="SC1A9.23"

CDS 22780..23601
/gene="SC1A9.23"
/note="SC1A9.23, possible gntR family transcriptional regulator, len: 273aa; similar to TR:O69996 (EMBL:AL022374) WhiH, sporulation transcription factor from Streptomyces coelicolor (295 aa) fasta scores; opt: 139, z-score: 267.9, E(): 1.2e-07, (29.8% identity in 292 aa overlap). Contains Pfam match to entry PF00392 gntR, Bacterial regulatory proteins, gntR family, score 30.50, E-value 7.1e-08."
/codon_start=1
/transl_table=11
/product="putative gntR family transcriptional regulator"
/protein_id="CAA22393.1"
/db_xref="GI:4007708"
/translation="MKHSAQGAARTGVGTGTGAAAEAVRIPVQPGAADRARGRAPGGA
EPDGPARGETHGEPVPRPRALVQRSSVRGQILDALRSALVTGELRPGEVYSAPVLG
ERFGVSATPVREAMQQLALEGAVEVVPNRGFRVLERGDRELAELAEVRALIEVPVWLR
LARTVPAEHWAE LRPLAEGTVRAASSGCPATYAEADRAFHRAALALAGNEQLVRIAGD
VHRRAQWPPAGSPSVRGRADLVADAHQHTALLDALIAGDLDDVVRALVGDHFTVTS"

misc feature 23008..23184
/gene="SC1A9.23"
/note="Pfam match to entry PF00392 gntR, Bacterial regulatory proteins, gntR family, score 30.50, E-value 7.1e-08"

gene complement(23847..26234)
/gene="SC1A9.24c"

CDS complement(23847..26234)
/gene="SC1A9.24c"
/note="SC1A9.24c, SecDF, protein-export membrane protein, len: 795aa; similar to many eg. TR:G3220156 (EMBL:AF024506) SecDF protein from Bacillus subtilis (737 aa) fasta scores; opt:802, z-score:1099.4, E():0, (30.7% identity in 740 aa overlap)."
/codon_start=1
/transl_table=11
/product="putative SecDF protein-export membrane protein"
/protein_id="CAA22394.1"
/db_xref="GI:4007709"
/translation="MKRSRPPSRSRSRTRSRSLNVRALVALAVMAGAVAIALTMPVRL
GLDLRGGTQIVLETKSTETTKADREATDRTVEVLGRIDALGVAEPTIVRSGENRVVV
ELPGVQDPKKAADVLRGAQLTVHSVLGAAEKPADATEGRTETAEDGERVLPDESGQS
LRLKAATLTGQDVKGADARFDQONGAGWTVTVDFKDSGSDRWAQVTGEAACNPAGDPT
RRVAIVLDDKIISSPQVDPSVSCGAGITGGSTQITGSFDDAEARELALLIKGGALPVP
VETIEQRTIGATLGDEAIDAGAAVIGTALTALFIIIVVYRLMGALATVALLCYGLIS
YAALAAVGATLTLPGLAGFVLAIGMAVDANVLVFERAREEQAAARTRPSTRSALTAGFR
SAFSAIADSNITTLIAAALLFFLASGPVKGFVTLGIGVLASMVSALVITRVLAEFAA
SRPAVFRPRITGISSTGPVRDALLRRDPFLMRRPRRWLAASLIVLVVAGSGILVRGL
NFGIEFTGGRLIEYSTATQVDPDRARDALADAGFPRAVVQSSGDGDLTVRTEELTDTE
AATVTKAVAE LGGETEKVRDELIGPSLGEELRRDALIALGLALAAQLAYLAVRFRLLF
GTAAVGALAHDVVILVGVFAWLGPIDGVFLAALLTVIGYSVNDSSVFLFDRIRELLGK
ERKAPFDRLTNDAILQTLPRTVNTGMGAVLILASLAILADDSLTDFALALLIGVGVGT
YSSVFTASPLAIELHNRDTGSRPGRRRRGRTRATGKSEKPGKQVPTSRTERQEV"

gene complement(26314..26703)
/gene="SC1A9.25c"

CDS complement(26314..26703)
/gene="SC1A9.25c"
/note="SC1A9.25c, probable secreted protein, len: 129aa; Contains a strong signal sequence."
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1A9.25c"

gene /protein_id="CAA22395.1"
/db_xref="GI:4007710"
/translation="MRRRLASVLTVLLTVLLPLVPAWPAAGAHVGPSPGLAAAAATAV
PHPALDLHADDGCTPVCAAQPRARHDQPAGRPTAPDQHPATTAHLGGGAAPCGHARTS
FAPGPVPVSPGRASHDSGRAPPVSSGT"
complement(26808..27590)
/gene="SC1A9.26c"
CDS complement(26808..27590)
/gene="SC1A9.26c"
/note="SC1A9.26c, probable luxR family response regulator,
len: 260aa; similar to a family of response regulators eg.
TR:O69961 (EMBL:AL022268) probable two-component
transcriptional regulator from Streptomyces coelicolor
(219 aa) fasta scores; opt:1079, z-score: 1434.4, E():0,
(77.5% identity in 218 aa overlap) and DEGU_BACBR
(EMBL:L15444) transcriptional activator protein DegU from
Bacillus brevis (236 aa) fasta scores; opt:378, z-score:
682.5, E(): 9.7e-31, (36.6% identity in 227 aa overlap).
Contains PS00622 Bacterial regulatory proteins, luxR
family signature, Pfam match to entry PF00196 GerE,
Bacterial regulatory proteins, luxR family, score 99.70,
E-value 5.6e-26, Pfam match to entry PF00072 response_reg,
Response regulator receiver domain, score 132.90, E-value
5.6e-36 and an helix-turn-helix motif from: 1 to: 260,
Score 1221 (+3.35 SD)."
/codon_start=1
/transl_table=11
/product="putative response regulator"
/protein_id="CAA22396.1"
/db_xref="GI:4007711"
/translation="MSDPSLPEPSEPSARGRHAGASTTPTTPTSASTTSAYGTPPASA
PSKIRILLADDHALVRRGVRLILDREPDLEVVAEAGDGAEAIDMARAEADLAVLDIA
MPRLTGLQAARELAALKPGLRILMLTMHDNEQYLFQALKSGACGYLKSVAADRDLVAA
CRAAMRDEPFLYPGAVTALIRNYLDRVRHGEETSDHILTPREEEVLKLVAEGHSSKEI
AEILFISIKTVQRHRANLLQKLGLRDLRLTRYAIRAGLIEP"
misc feature complement(26874..26957)
/gene="SC1A9.26c"
/note="PS00622 Bacterial regulatory proteins, luxR family
signature."
misc feature complement(26952..27293)
/gene="SC1A9.26c"
/note="Pfam match to entry PF00072 response_reg, Response
regulator receiver domain, score 132.90, E-value 5.6e-36"
misc feature complement(27393..27590)
/gene="SC1A9.26c"
/note="Pfam match to entry PF00196 GerE, Bacterial
regulatory proteins, luxR family, score 99.70, E-value
5.6e-26"
gene complement(27599..28510)
/gene="SC1A9.27c"
CDS complement(27599..28510)
/gene="SC1A9.27c"
/note="SC1A9.27c, probable sensor kinase, len: 303aa;
similar to a family of sensor kinases eg. TR:O69960
(EMBL:AL022268) probable two-component sensor protein from
Streptomyces coelicolor (358 aa) fasta scores; opt:1037,
z-score: 1477.8, E():0, (59.2% identity in 299 aa overlap)
and TR:Q52558 (EMBL:U02041) a two-component sensor
regulating virulence genes of Pseudomonas solanacearum
(502 aa) fasta scores; opt: 388, z-score: 475.5,
E():3.3e-19, (31.5% identity in 270 aa overlap). Contains
hydrophobic N-terminus."
/codon_start=1
/transl_table=11
/product="putative sensor kinase"

gene /protein_id="CAA22397.1"
/db_xref="GI:4007712"
/translation="MVLGFATALLLWAPVTVSVPIILLTEAVVLVVGMGVMLVANGALL
RWGLAPLERLTRLMTTVDLLRPGQRLPISGGGEVPELIRTFNAMLDRLENERATSSAR
VLLAQEAERRRIAQELHDEVGQSMTAILLVLGRAADDAEELRDELHQAQEITRESLD
EVRRLLVRRLRPGVLDLGLISALSSLTHDFATHTGLRVVRRFDADLPVLDHETELVLY
RVAQESLTNAARHADAERLEVGLAHADAATVLTVADDGRGIEAAHEGAGIRGMRERAL
LIGAALDITSAPGAGTRIRLTAPLPRK"
28890..29216
/gene="SC1A9.28c"
CDS 28890..29216
/gene="SC1A9.28c"
/note="SC1A9.28c, unknown, len: 192aa; some similarity to
TR:066611 (EMBL:AE000680) putative protein from Aquifex
aeolicus (145 aa) fasta scores; opt:188, z-score: 294.0,
E():4.4e-09, (36.6% identity in 93 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1A9.28c"
/protein_id="CAA22398.1"
/db_xref="GI:4007713"
/translation="MSLDASRIEPRPERLTAHEARQRLHARNTRVTQLQALAESGQA
DDQLMSAQKAAIERVLKEIDEAFARVEEGTYGACLGCGKPVGERLEILPYTRYCVAC
QRRAAA"
gene 29286..29400
/gene="SC1A9.29c"
CDS 29286..>29400
/gene="SC1A9.29c"
/note="SC1A9.29c, partial CDS, unknown, len: 38aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1A9.29c"
/protein_id="CAA22399.1"
/db_xref="GI:4007714"
/translation="MNNQIIGDRDTRLPLSPEDLAALRDNLREQRLFREEQ"

BASE COUNT 4022 a 11280 c 10476 g 3622 t
ORIGIN

```
1 gatcacctcg tccacgtcgg cggccagcgc gaaggcctcg tccagcggct gcccggcgcg
61 ggtgcggtcc ccggccaccg tgtacgacag ccaggcgggc acgtccagtc cgcgcaccgc
121 cctcagcagc gccgccgcct cgtcgggtgtc cgggaccgtc tccagcgcga ggacgtcggg
181 ccgtgcggcg gccagcacct ccaggcgggg gcggtggaag cgctccagtg cgcctcgtcc
241 gaggccgtac cgtccccggg actcgggaacc gtccgcgagc atcgctccgt acgggcccgc
301 cgacgcgcgc acccacagtg cgcgctccgg cggggccgtc cgcgcgccgc gggcgccctc
361 gcgggccgac gccacgctga gcgcgagcag ttcggcgggc cgctcccgcc cgatcccgcg
421 ccgggcgaag ccctcgaagg tggcctggta gctggaggtg atcgccacct cggcgcccgc
481 ctcgaagtag gcgaggtggg cccgcgtgat cgctcggggg tctcggcgga gcagccgggc
541 cgaccacagc gcgtcgccca ggtcgtgccc ggccgcctcc agctggttgg acagcccgcc
601 gtccagcacg agcgggcccg aggcgagggc gtcggcgaa gtcgctggtca tggcacgagg
661 ctacgacctg tccggcgggg ccacgggccc ttgtcccccg agcgagctac agacaggtgt
721 ccaccgtgag gtgacgatct ccggccgccc gatccgtagc gtccgggtga gcccgcgcca
781 ctcgagccac gggcccgcag aaggagttcc cgtgaggttc gtgtggcagt tcttggccgt
841 gctggtggcc tacgccatcg gcggcatcgc cgtccaggcg gtgaaggaca acgactggct
901 cacgctggtg gtcgggtctca cgtcgttggc gctggtggtg ttcgtgtacg cctgggtggt
961 gcggcgggacc gagcgccggg aggcctggga cgtggccctg gacggggccg cggccaaggc
1021 gggctggggg acgctgatcg gcttcggact gttcggggcc gtcattacga acctcttcgc
1081 ctccgggtac tacgaggtcg acggactcgg ctccgtccag ggtgcgatcg ggctggtcgg
1141 gttcatggcc gccgccgccc cgacggagga ggtcgtgttc cgcgggggtc tgttcgggat
1201 catcgaggag cacatcgcca cctacctcgc gctgggcctg accggcctcg tgttcggcct
1261 catgcacctg ctcaacgagg acgccaccct gtggggcgcc ctcgccatcg ccacgagggc
1321 cggttcatg ctgcgcgccc cgtacgccgc caccgcgaac ctgtggctga cgatcggcgt
1381 gcacttcggc tggaaacttc ccgcggcgcg cgtcttcagc accgtcgtct ccggcaacgg
1441 cgacagcgag ggcctgctgg acgcggccat gtcgggtccg aagctgctca ccggcggcga
1501 cttcgggccc gagggcagcg tgtactcggt gggcttcggg gtctcgtga ccctggtgtt
1561 cctgtggctg gcgcaccggc gcgggaacat cgtggccttc ggctcgcggc gccgtgccgc
1621 gggcgccaac tccgccgcta cacttccccg gtgatcgatc gacgacgggt cctggagctg
```

```
1681 tggcgctcgac tcgacgtcac ggtccggggac ctcccgcctcg ggggtgctgct gctcgctcgcg
1741 tcgctgctgc cgtcgctccg cggtcagggc acggagatcg gcggcctgcc gaccgcgcgc
1801 gccgacgcgc tggccggggt ggcggccgtc ctccagtgca tccccctggc cgtgcgccgg
1861 cgggtggacgc tcctctgcct caccctggtc tcgctcggtc tcgccctcga ccaactgcgc
1921 gcctaccacc tgttcgcggg cgcgcgcgtc ccgatcggtc tgatcaacgc gggctccac
1981 caggagaagt accggcgtgc cacgcaggtc accgccacc tcgggtacgt ggccatggcc
2041 gtcgggctga acgcgcgcgc cggcgacgag acgctgggtc agtacgtgac gttctatctg
2101 gtccctcgcc tggcctgggg catcggcgcg tggatgcgtc ccgcgcgggc cgcggaggcc
2161 gaacgccgca gccgggtcgc cgaggacgcc cgcaacgcgc aacggaccgc catcgcccgc
2221 gagttgcacg acgtcgtgac ccaccacgtg acggcgatgg tcgtgcagtc cgaggcgccc
2281 cggatccgga ccgcgcgcgc cgacgcgtc gacgagagtc tggccgcggg cagcgacacc
2341 ggccggcggc ccatcaccga cctgcggcac ctgctcgacc tcctcaacc cgaccacggc
2401 accgccgagc ccaggacacc acccgtcggc cgggtgctca cgctggtcga gcagaccgc
2461 cgggccgggc agccggtgga gttcaccgag gagggcacc cggcggcggc caccggcagc
2521 tccgacctcg tggcctaccg tgtcgtccag gaggccctga ccaacgccct caagtacgac
2581 cacggcgcca ggacctcggg cctgggtgcg cacggggaac gggagatcac ggtggaggtc
2641 ggcacggacg gctccggctc gggggccgcg tcccccggcg gcagcgggcg ggggctggcg
2701 ggcccgcggc agcgggtcga cgtactgggc ggcgagttca gcaccgaccg tccggcgga
2761 ggccgcttcg tggtcggggc ccggataacc ggccgggagc ggggaagcac ggcattgagc
2821 cggcgatccg ggtcgtgac tcgcgacgac aggcactcat ccgcaccggg ctcgcgacga
2881 tcgtcgacgc ccagcccgac ctcgaggtgg tcggcgagtg cggggacggg cagaccgggg
2941 tcgacctggc ccgcgaactg cggccggagc tcgtggtgat ggacatccgc atgccggtgc
3001 tcgacggcct ggaggccacc cgcctgctcg ccggtgccgc agtggcgcac cccgtcaagg
3061 tgctcggtg gacgacgttc aacctggacg agtacgtcta cgaggcgttg cgcgcgggcg
3121 ccagcgctt cctgctcaag gacgcgcgcc cggaccggct gctgcacggc atccgcaccg
3181 tggccatggg cgcggcgctg ctgcacccc atgtgacgc ccgcctcgtg ggcgctacg
3241 cggcccgat ccgcccgcgc gagggcaccg cgcgggacat tccgctgacc ccccgggaga
3301 cggaggtcct gcgcctgac gcgacgggc tgtccaacag cgagatcgcc gcggccctcg
3361 tgatcagccc cgagaccgtc aagaccttc tctcccgcat cctcaccaa ctcgacctgc
3421 gcgaccgcgt ccaggcggtc gtcttcgct accgccacgg gctggtgacc tgacgcccgc
3481 ccgcccggcg ctacgcccgc caaccggccg cgcattgtgc acagacgccg gtcgatgcgc
3541 aggttgtacc aggccagcca gccaggaac acgaccccga ggccgagcat caggcttccg
3601 gcaccgaccg agccggagag gtaccacagg gccgaggtgc cgaagaagat gacggccagc
3661 atcgggaacg cccaccagcg gttgcgcgc agccgtcct ggcgcgagtc gacgaacgcc
3721 gccatcgccc gccgcgcgtc cgggtcccgc ggccgaggcc cgccttgag gatctgccg
3781 tccatggcgg gacgtcgtc cggttccgca ccgaccgcc tggcgtcgcg cacgcgcgcg
3841 cggcgacacc accacacgat ccagacggcg gccggcaccg cgcaccagc agccgacagc
3901 cacgccgacg tgcggttcag gaccaaccgg accacgacgc cggccagcag ggccatcact
3961 ccgatgccc acaggtgtc ccgcctgcgc agccgcggac ggtcgtcgct tctctctgtg
4021 ctgcctatgt cacctctcat ggggaagacg tggacgtacg gtgcagaccg gataccgcgc
4081 ttcgcccgga acacgtcgcc gggacacttc ggccggcgcc gaagcaaccg gcccatagcg
4141 tgcgggacat gaacgacacg cgaaccccca cgagcacctg caccagcacc cgcacgcccg
4201 ccgacgccgt caccggcatg gtcgaccacg tccgtcacct ggccgcgggc tggaccgct
4261 gggacgggac gcccgcccac gtcgacggcc gcgtctacac cccgcacaag cgcattcggc
4321 gggtcgccga ccacctggtc gaccacttcg cggagctgga ggcccggtg gcgggggagg
4381 agacgcagcc cgaccactgg cagcctccc tggtcaccac ggaagcggac cgcgcgcctc
4441 tcaccgcca ggatctcgac gaggcccgca gccgcctcac ccggtcgcc cggatctggg
4501 ccaaccgct cgacgcgtc accgacgagc agctcgacca ctgcgcgggc gagggctgga
4561 gcttttcgca actggccgcg cactcgcggc agtcgggta ctacgccgac gcggtgggcg
4621 acctgtcgtg acagcgaccg ccttcgcgc cgtgcaccgc gccggtgagc cctgctgct
4681 gccgtgcgc tgggaccacg cctcgcctt cgcctcgcc gggcagggt tccgcgcgt
4741 cggcacgacc agcctcggtg tcgcccgcgc ggccggactg cccgacgggc cgtccgcgac
4801 ccgggacgag acgctgcggc tggccctcgt cctcggtcg gcccccttc cgtctcctg
4861 cgacgcggag gacgggttca gcgacgacc cgacgaggtg ggggagttcg cgcggcagct
4921 ggccgcgtc gccgcgcgtc ggatcaacct ggaggacggg ctcgggccc cggccggga
4981 cgccgcgaag atcgccgcgc tccgtcggc cgcgcccggc ctgttcgtca acgcccgcac
5041 ggacacctac tggtcgggcg acggcgacgt cacggagacg ctgcggcgcc tggaggccta
5101 ccgcgaggcc ggtgcgcagc gcgtcttcgt ccccgggctc accgacccc cccgcattcg
5161 gtccctcgcc gccgcgttc acgtgccact caacgtcctg tacacgccc cggaccgcg
5221 cctcgcgcac ctgcgcgacc tcggcgtagc ccgcgtcagc ctcggtcgc tgcctaccg
5281 gcgggcgctg ggcgcgcgc tcggggcggc cgcgcagctc cgggcgggc cggaccggg
5341 cggtcgacg cccacgtacg acgaggtgcg ggcccggggc tgacctcagg cgcgcggcca
5401 gttctccagc gccacgtgca gcgcgtcgac gccttggtc caggacgcc gcacgtccc
5461 cggggcgccg aaggcgccgc aggactccag cgcgcagtac ccgtggaagg tgcgtcgcg
5521 caggcgcacg gcgtcgggtg ggtcgggctc ggtgaggccg taggaccgca gcatgccgta
```

5581 ggtgatctctg gcggtgcggc gcagcgcggc ggagtcggcg accagggact ggtcgacgcg
5641 gatctgcgtg gcggcgtagc ggccgggggtg gcgcaggcg tacgcccgt aggcgcccgc
5701 gaaggcggcc agcgccctct tcccggcccc gcgggccacc gctccggcga tctctcgat
5761 cagctcgccg ccggccagca gcgcgaccct gacgcgcagg tcctggaggt tgcgcacgtg
5821 cgtgtacagg ctggcgctct tacccccga gtggcgggcg agcgcggaga cggtgaccga
5881 ctcgaaagccc gtctcgctcg ccaggtcggc ggccgcggcg acgacgcggt ccgtggtgag
5941 acccactcgg ggcatgggccc gccttcccgc aacctggagt tcctaggttt catgctaggg
6001 caggacgggc ggcagccgcg gatcaccgc gagtcgttca acgttcacct tccccccga
6061 ttctctctga agccgcgcac ggcgcggtc accaccacgg acatacgac gaccacgagc acgccaccga
6121 cggaacccc caccgtcacc accaccacgg acatacgac gaccacgagc acgccaccga
6181 acaggtcctg cccgcgcctc tcgacaccag cgtccccgac gaggccctga gccccgcga
6241 gcagtcgcgc cggggcctgc tgcgcctgc cggcctgctc ggcgccggac tggccgcgg
6301 taccgtctct gccccggccg ccaccgccac cgccgcccc gcgcggggcg cgagcaaccg
6361 ccgcgcggg aagggcttcc tctggctggc cggcgaccac cacatccaca cccagtacag
6421 cagcgacggc aagtaccgcg tcgtcgacca ggtccgccag ggcgcccggc acggcatgga
6481 ctggctggtc atcaccgacc acggcagcaa caccacgcg aagatcggt tggagaagg
6541 caaccccgac atccgcgagg ccgcgcgcg ccacgaggac accctcgtct tccagggcct
6601 ggagtggaac atcccggccg ccgagcacgg caccgtcttc gtccacccc gcaagcacga
6661 ggtctccgtc ctcaagcagt tcgagaccga ctacgacggc agcgtcaagg gcgcgggga
6721 ctccaccccc gccaacgagg ccctcgccat cgccggactc tccttctctg ccgaccagg
6781 caagcggcgc aaggtcaagg acgcccgtgat gctcgccaac caccggcgcc gcaagggcat
6841 cgactcccc caccgagatcc gcgcctggcg cgacgccacc tcccgcggcc accagatcgc
6901 ggtcggcttc gagggcgccc cgggccacca ggcgggcggc ctgcccgagc ccctcgcccc
6961 gggcgggccc cgcgccatct acgacggcag cccgagcgcc aactccttcg ccggtaccct
7021 cctggagagc taccgcacct gggcgggctt cgactggatg accgccaccg tcggcgccct
7081 gtgggacagc ctcatcgccg agggccgccc ctggtggatc accgccaaact ccgactccca
7141 ccaggtctac gccgacaccg gcgcccgcgg cgcgggcgac ttcaacgcca acggccgcta
7201 cgacgacccg gtctacgcgg ggcagatcga catcaccag aacgacttct ggcccgcca
7261 gtacagccgc acccacgtcg gctccgacgg cttctcttac gccgcctca tggacggcat
7321 ccgggcccgc cgctctggg tcgaccacgg gcagctcgtc agcgccctcg acgtccgct
7381 ctccggcggc ggccgctggg ccacgctcgg cgcgccctg cactccgca ggggacccc
7441 ggtcaccttc tccatcgacg tcgcccggc cgcgggcccc aactgggccc gatctgtg
7501 gaagctggcc cgctcgacg tcattccagg cgacgtgacc ggaccggccg ccgacaagg
7561 caccttcacc gcgcgcaccg ccgggtcgtc caagtctac gaggtcgaca aggaaccgg
7621 caccgtccgc ctacactacg acctcgccg cgtcgaccgc ccgctctacc tccgacccc
7681 cggcaccgac ggcaaccggt ccgcgctcgg ctcgctgggc gccaaagtcg acccgccgg
7741 ccccgccatc gacgtcgtcg gcgacgccga cccgtggcg gacctgtggt tctactccaa
7801 cccggtctgg gtctgcctc cgtgacccc cccgcgcgc cgtacgtcct cggcatcgac
7861 accgacgcga cgaccctgcg cgaggccgat cactgtctc aggcgtggtg ggccgaactc
7921 gacctcccc agggcgtctt cggctgcacc cactgtgtc gggacggccg cccgcgggtc
7981 gccctgtccc tcgcccgcga ggcgagccg gtctgcga cccgcggga cggctcagc
8041 ccccggggcc acgaggtccg cgacggcacc tgggacgaga gcggccgcgc ggtcctcttc
8101 cccggtcggg ccgcccacac cggcaccctg accctcgccg agctgtctgc ccgctccgc
8161 gtcgaccggg tgaccgtcct cggcacaccg gacgagccga gcccgacac cgggtggtc
8221 acccggaacc acgtccgccc gcactggcag gacggccgc tcgtgctggc cgccatgccc
8281 gcggtcggcg gcacctcgt accgttcgag gaccccgacc cgacaccctg ctgcgcgac
8341 cactgacccg gcccggcgccc gagggcgcg ggggtgctcag cccgcgagca cccgcgcac
8401 gaacgcgttg gtgaacttcc ccgcccgggtc cagcgcgccg gccagcgcgc cgaagtccgg
8461 cagccgcggg tacagcgcac gcagctcgcc cgccgggacg gtgaacacct tccccagt
8521 cggacgggcc gcgaagggga cgagcgctc ctccagccgc cgcaccaccg cgaccaccg
8581 cgccgtgtcc tcgaccaggg tgaagtgcgc ggccacgggt tcccggcgt acgcgggct
8641 cagccactgc cgtcggcgg cgaccgtgc gatctcgag gtctggagca cggcgcgag
8701 cgtctccgt atcgctcca tcgctgcag ggcgccagg gcgtgctccc ggggcatcag
8761 gtactccgac tgcaactcgg caccgtgct gggcgtgaac tcgcgcgga agtgcgag
8821 ccgctcgtgc cagggcccc gcaccccga ctgctccgtg cagttaccg cgggcatgcc
8881 cggcaccgga tgcatttct cggcgccgg ggcgcgtac gggaaacct ggcggccag
8941 gtcggtgcgc cgcttcagcc acacctgcg gaaaccgggc gcccgccagt cgggtgaacg
9001 actcacgtg tacgcccgc ccattaccgt ctcgaaactc gccgggtcca acccgccag
9061 cggcagctcg gtgaagacgt gctgttccat ctcgtagcc ggctccagg cgagtgtgag
9121 cgagctcacc acgcccagc cgccgacgga ggtcaccgccc cgccgaacc gctcgtcgcc
9181 ccgcgcgac accacggctg aaccgtccgc ggtgaccagc tccacctcgc gcaccaccga
9241 ggccagcgaa ccgttgccca ccccgaaacc gtgggtgccc gtggccaccg acccgcgac
9301 cgagatgtgc ggcagcgagg ccatgttcgg cagcgccagg cccgcgcgt gcaccaccg
9361 ggccagctcc gcgtaccgca caccgcgcgc gaccgcacc gtacggggcg ccgtgtccac
9421 gtccaccacg gacggcaggc ccgcccagca cagcaggaca ccccgctcgc ccggtcggc

9481 gatctcgttg aaggagtgcc cgctgccag caccgcacc ctggcgctgt ccgccaccag
9541 ggcccgccag gcgtccagg agtgccgag cagcagttcc ttccgctgt acgtgatgtt
9601 gccggccag ttggtgacc tgatgtcgt catagcgagg aaacctacc ggcgctctg
9661 gagggcccc gcgcgggaag cctctccgc gggggcacc cgtgagaagt cgaacgccg
9721 acccgggagg agacacggga tggacagcc taccgctgt gtcgaggatc tgatggagcg
9781 gttcccgcac gtaccgcgc aggcgctct caaggaggac ctgctgcgc gcggtgtcgc
9841 cttcgacccc tccgcactca gtgataccac caacgaggca gccggcgagg tcaagccgaa
9901 gtcgtacttc atcttctcct tcgaccacgc caccctgccc gagttgggag aggcgcctt
9961 gcgcccgcgc cccgaggaga tcctctcac cggcgcccc tacgacctgc gccgcaccgt
10021 cgtctcggtg cgcgtgaacc cggcctcgcc ctaccgggtc gccgcgaacg aggacggcgt
10081 cctcggcctc tacctcgacg gcaagcggat cgcgcagctc ggctgcccga ccatgcccga
10141 gtactaccgg cacaagctct ccaacgggaa gtcggtcatg gaggtggccc cgaccatcca
10201 gtggggctac ctgatctacc tgaccgcgt ccgctctgc cagtacttc ggcaccaagg
10261 ggagtgccag tactgcgaca tcaaccacaa ctggcgccag cacaaggcgg ccggcgggcc
10321 gtacaccggg gtgaaggacg tcgacgaggt cctcgaagcg ctggagatca tcgacaagta
10381 cgacaccgcc aagatctcca ccgcgtacac gtcaccggc ggcgcgatca cgtcgaagg
10441 ccagggtctg gacgaggccg acttctacgc ccggtacgcc aaggccatcg aggagcactt
10501 ccccgccgcg tggatcgga aggtcgtgc ccaggccctg cccaagccgg acgtccagcg
10561 cttcaaggac tacggcgtgc agatctacca cccaacttc gaggtgtggg acgagtacct
10621 gttcaagatg tactgcccc gcaaggagcg ctacgtcggc cgcgacgagt ggcaaacg
10681 catcctggac tccaccgagg tcttcggcg ccgcaacgct atcccaact tcgtcgggg
10741 cgtcgagatg gccgagccct tcggtctcaa gaccgtcgac gaggccatcg agtccaccac
10801 cgagggcctg cgcttcttca tgtcgacagg catcacgccc cgcttcacca cctggtgccc
10861 cgagcccacc accccgctcg gcaagaccaa ccccgacggc gctccgctgt agtaccatc
10921 ccgcctgctc caggcctacc gccagaccat ggaggactac ggtctctcct gccccccgg
10981 ctacggcccg cccggcgag gcaacgcgt cttctcggtc agctccttca tggacagcct
11041 gccggaaggc gcgcccgtgg aggtgtgagc ggacggttcc ggggtttatt cgggtggcgc
11101 gccttcgggc gcgcgactag ggtgaccgat gtcgagcgg cggcccaagc ggccgcgac
11161 aggggtcttt tccgaggagg tgtcgaccaa tggctgtcac tgtgacgggc gctgcccga
11221 tccaggagtc catcgagtcc acctccgagg ccaccggctg atctcagccg ctccctgagt
11281 tctcgcgccg gggccgcacc ccgaagggtc accccttgac ttccacgtct tccactcca
11341 accactccgc cctgtctctc tacggctggg acgactcctg ggccgacgag ttccgcccgt
11401 acgcccgcga agggctgctg cccggccggc tcgtgagggt cgaccgcgga cagtgcgacg
11461 tcgtcaccgc ggacggcgctc ctgcgcgctg acaccgcgtt cgtcacgccc caccaccgc
11521 tgcgctcgt ctgcaccggc gactgggtcg ccgtcgaacc gggcggaac ccgcgtacg
11581 tacgcacgta tctgccccgc cgtaccgcct tcgtgcgtc cactcgtcc aagcgtccg
11641 aggggcagat cctcgcggcc aacgtcgacc acgcccgtgt cgcctgtcc ctgcggctg
11701 agctggacct cgcccgatc gaacggttcc tcgcccctgc ctgggaatcc gggcgcgagc
11761 ccctggctgt cctcaccagg gccgacctc tgcccgaccc ggtgacctg gcgtacctg
11821 tccaggagct ggagaccgcc gcgcccggcg tgccggtgt gccggtcagc gccgagcagg
11881 gcgaggggct cgacgtgctc gccgcgctg tctccggcgg taccgcccgt taccgcccgt
11941 agtccggcgc cggcaagtcc acctcgccca acgcgtgct cggcgaggcc gccatggagc
12001 tgcagccat ccgcaacggt gacggcagc gccggcacac caccaccacc cgcaacctc
12061 tcgcccgtgc cggcgggggc gtgctgatcg acacccccgg actgcggggc gtcgggctct
12121 tcgacgcggg caacggggtg gaccaggtgt tcgccgagat cgcggagctg gccgaggagt
12181 gccggttcca cgactgcgcc cagcagagcg agccgggctg tgccgttctc gccgccatcg
12241 acagcgggtg actgcccag cggcggtg agagctaccg caagctgatg cgggagaacc
12301 agcgcacgt cgccaagacc gacgcccggg cgcggggcga gatccgcaag gactacaagc
12361 gccggggcgc catcggaag gcggcgatgg aggccaagcg ggggtggcctg aggtagcccg
12421 ccggccgctc caacgcgatg tccgatttcc gccagcccgg tccgcccgtc tggcgagac
12481 tggacagcgt gacaccacag acagtgcgc cggcagagca cgcggacgag cgcgaggacg
12541 tccggtacga ggagtgcg agccgggacg cccggttcga cggcgcttc ttcttcgccc
12601 tcgagacgac cggcatctac tgccggccca gctgccccgc cgtcacaccg aagcggcgca
12661 acgtgcgggt cttcgcgacg gccgcgcgc cgcagggtc gggcttccgg gcctgccggc
12721 ggtgccgccc cgacgccgtg ccgggctccg ccgactggaa cgtccgcgcc gacgtcgtcg
12781 gccgggcat gcggctgat gccgacggcg tgggtggacc cagggcgct gcccgactcg
12841 ccggacggct cggctacagc gcccgccagg tgcagcgcca gctcaccgac gaggtgggag
12901 ccggaccgct ggccctcgcc cgcgcccagc gggcgccac cgcgcgggtc ctgctgcaga
12961 ccaccgtcct gccggtcacc gagatgcct tcgctccgg gttcgccagc gtgcggcagt
13021 tcaacgacac gatccgggcc gtgtacgcgg ccaccccgag cgagctgcgc gccgcgcgcc
13081 cggcccgcga ccgggcccgc cggcgacagg ccacccctc cgcggcgctc ccgtgcgcc
13141 tggcccaccg cggcccctac caggccggcc ccgtcttcga cctgctccag cgggagggcg
13201 tcaccggcgt cgaggaggtg agcggcgaga ccggccgccc cctctaccgg cgcacctgc
13261 gcctgccgta cggcaccggg atcgtcgccg tccaggaacg gccgggcccg gcgggcaccg
13321 gcagcggggg ctggctggag gcccgccctc acctcaccga cctgcgcgac ctgaccacct

13381 ccgtacagcg gctgcgggcg ctgttcgacc tcgacgccga cccgtacgcc gtcgacgagc
13441 ggctcggcgc cgaccccccg ctgccccccc tggctcgccg ccgccccggg ctgcgctcac
13501 cgggcaccgc cgaccccccg gaactggcgg tgcgcgccct ggtggggcgc accgaggccg
13561 agcggctcgt ccagcggtag ggcaaggcgc tcgacgcgcc ctgcggcacc ctaccccacc
13621 tcttccccga accggacgtc ctgccccggg ccgcgccgca cggcaccocg ggcgcgctcg
13681 ccgcccgcct cgccgacggc gccgtacgcc tggacccccg cgccgaccgg gacgacgccg
13741 agcgcgcgct gtcgcgcgtg cccggcctgg acgcccgtac cgtcgccgtc gtgcgacccc
13801 gcgcctcctg cgacccggac gtcgccccgc ccggcgccgc cgtccccgac acctggcgcc
13861 cctggcgctc gtacgcactg aaccacctgc gcgcagcagg agagtgggag aacgaccgat
13921 gaccacgacc acgcccgacg ccacgaccac gagcatcccc gccgagacgt actggcacga
13981 ggtcgacagc cccgtcgggc cgctgctcct caccgcgcgc tccgacggcg cgctcacctc
14041 cctgtccgtg cccgggcaga agggcgcccg gagcgtccgg gacgggtggc ggcacgacgc
14101 cgggcccttc cgggtggccg aggaacagct cggcgccctac ttcgcccggg agctgacgga
14161 gttctcgctg ccgctgcgcg cgcagggcac cgccttccgg gagcgcgtgt gggcggcatt
14221 ggacgacgtg ccttacggcg cgaccacgac gtacggcgag atcgccgcgc gcatcggggc
14281 gtcgcggccc gcggtacggc ccgtcggggg tgcgatcggc gcgaatccgc tgcgtatcct
14341 gcgcccctgc caccgggtga tcggcgccga cggctccctg accgggtacg cggcgggcgt
14401 ggaacgcaag acgcggtgc tgcgctgga gggcgccccg ctacgccggc cagtaccgct
14461 cccagccacg ccccggtgat cagcaggcag gcgaacagct cggtcagcag gtcggcgccg
14521 ccctggcgca tcgccgtccg cagcgacgcc atcgccctgc cgtgccgcc cagccgcagc
14581 cgctcggcga gtagatgcc gcccatgaag ccggggaccg cgcgcagcac ggggaccagg
14641 acgaagccga ggacggcgcc cgcgcccgcg tacgcccgtc gccggcgcat ggcctcgctc
14701 cggcgcttcc tgcgggtcgg cagcgcccag cgcaccgccc gggacagcag cagcacgacg
14761 gtggaaccga ccaggaccgc ccacgcgagc ggcgcgggt ccttcagcgc ccaccacagc
14821 accccggccc acaccagcca cgatcccggc acgcccggca gcaggacgac gcacggcccg
14881 agcagaatga ccaggccggc cagcaggagg tcccacgctc ccatctgtcc aggggtcagg
14941 agtccggcca ggtccgcaga agtccggccc ggcgggcggg gccctcagt cctcctggc
15001 caccagctcc cgctcgtagg cgtgccagcc cagctggagc cgggtggtga ccccggtcag
15061 ctccatcagg cgtttcaccg gtcgctgcac ggtcctgagg cccaggtcga gctgcttggc
15121 cacgctcgcg tcggtcagcc cggcgagcag caggagagg atctccagat cggtgccgtc
15181 ggggccgtcc ggggcctgct cggtcacgcc ggaggcacc agacgcagcg ggagggcgctc
15241 ccgccacacg gcctcgaaca ggccacacag cagttcgagc agaccgctgg cgtgcaccac
15301 cagcgcgggc ggttcagagg agcgcgcggt gagcgggacg agggcgagg agcgctcggc
15361 cacgatcagc ttggtcgga cccggtccac gaccgcgacc tgcctcgcgc ggcccagcgc
15421 ggcggtgacg tcggtgatgc cggtcggcag gtcgaggacc gaccgctcga cgaccaccgc
15481 gtagcggaca ccgcgcccgg tggcctgctc ctccgcgtcg ttctccatgc cggtcacggc
15541 caccggcttg tcggtgacca gcgcgcacac ctctcgtc gcgcccagct ggagctggag
15601 gaagcgctgg gcgacggccc cggcgccggt caccacctcc accaggtcgt gcaccgcggg
15661 ctcgcccgcc gccgcccggc actcctccgc cagcagcgcc gccgccagct ccgcccgcctc
15721 cagctcgtgc cgctgctggg tgagcagggc gccagcgcg acgcccggcg gggccgcgac
15781 ccagcgtccc gggcgggccg aggactgggc ggcgagccc ttctgtcca gccggcgag
15841 cgcgcgctcg gtgtcccgtt cgcgcgagcg cagccgcctg gccagatccg gtacgtcggc
15901 cgcgcccacc gacaccagcg cccggtacgc cgctcgtgc gtgtcctcca gcccagcac
15961 tcccagcatc aggtgtcgcc cctcccaga accatgccca cctccacctg cgcggacgtc
16021 cgtggcggga aacggccacg gcgcaaacc gcctcgggac atcatcgccg taccgcccgt
16081 cactctgcca aggtggcgtc gccgggtggt cctcccgtgg ggggtggggc cgcccggcca
16141 cccacctcga gagaactgtt cgacgcgtcc cccgacacgc cgccgggcgc acgcccgtcc
16201 gcccgaatt tgccgatgcc ccgtttccat ccggcccat cggtggaaca tcaggagcat
16261 gagtcagcag gggggaaggc ccaccggtcc cgccggtccc accggacctg aggacgactg
16321 gtgggggagc ctgtacgacg actccaccga cgacaccggc cccacggccg caccgatcc
16381 cctggacgac cgcttcgcct cggcgctcga cactaggcc cgcggaccgg cgcccccg
16441 ccccgcgag cccacgcccc cggggccgce accgtcccc accgcgccg aaccgggacg
16501 cccctggtgg cccaggaccc cgcgcgcgc cgtccccggc cccgcaccg gggacgacgc
16561 cccaccgctc cacgcggacc tgcgggacc ggacgacggc gaccgaggc acgacgcccc
16621 accgcccgc gccgggggga ccggctcggg gagacgcggg ccgtaccggt ctccgggctc
16681 cccggaggcg ggggacggcg ggcccggcga ggggtggccc ccgcgggctg ccgggcccgc
16741 ccggcccacg gctccctggg agcctccgce cgcgaccccc gcgggcccga cgcggcccgc
16801 ggccgtgagc gccgaccccg gcgagggcgc gtcggacgcg ccgcccggcg cgcccgcggg
16861 aactccgccc ccgggtgaag cgtcgggtcc cgtgtcgtcg ccggcgctcg gtcggggcg
16921 gcctgccgag ccgaggcggc cggatccgcg tgggcccacc acggccgggt cggggcggc
16981 cgacctcaga gacgcgccg aggcggtccc tccgcccgg accctgccc acctgccgc
17041 gcttccgccc cgcttccctc cctcgacgga ggccgaccgc gactacgtcg gcagcgggcc
17101 gccacctac gacgcgcgag ccaccgcgct gccgcccgc gatcccgac gcctcgacgg
17161 cctggtgccg gacaccgtcc tggaggcgcc ccggtacggg acgtgcacgc tgcgggccc
17221 ctccgtgcgc ggggactccg cgcgctaccg gggcgagccg cgtcgcgacg cgctgctcgt

17281 cgcccggttc ggcgcggag agcaggcgct ggtcctggtg gcgatggcga ccggcgccccg
17341 tggcaccgcc gggcgcgacc gggcgcgccg cgaggtgtgc cgggtggatcg gacggggcgt
17401 gggcgcgacc caccgcccgc tcgcccaggga cctgcggggc gcccgcgcg gcgacctgaa
17461 gtccggcctg caccgacctca ccgaccgcag cctcgggccg ctccggggccg gtgctgccga
17521 acagggcctg gcccccgacg agtacgcggc caccctgcgc tgccctgctgc tggccgcccga
17581 ccccggggtg cgtacgcggg tgttcttcgg cgtcgggcg gcgggactgc tcagactgcg
17641 cgacggcgcc tggcaggaca tggagccgga cgccggtgac gtcacgggag agcccgtcct
17701 cggtctcggc tccgcacccc gcgacgcacc ccgcgacgca ccgcgcgaga cggccgagga
17761 cgaccggctg accatggacc tcggcatcac cacatccccg ggcccgtacg agggcgccgc
17821 cgccgggccc ccccgcgagc ccttcgggtt ccgggcccct ccgtagcccgcc cgggtgatgt
17881 gctcctgatg tgcaccgcgg gcctcgccga gccactgctg agcgagcccg gactgggcga
17941 actcctcgcc cggcgatggg ccgcccgcgc ggcccccga ccgggggagt tcctcgccga
18001 cagcggcgtg cgggtcaagg gatacgcgga cgaccgtacc gcccgcccg tttgggaggc
18061 gtgagaccgc gcgctgtggc atagctggac ccatggccaa acagaacgtc gcggaacagt
18121 tcgtcgacat cctcaccgcg gccggggtcg agcgctcta cggagtctgc ggggacagcc
18181 tcaaccccgt cgtggacgcc gtgcgcggcc attccggcat cgaatgggtg cagctacggc
18241 acgaggagac cgccgccttc gcccgcggg cggaggcgca gatcaccggg aagtcaccg
18301 cgtgcgcgg ctctcgcggc cccggcaacc tccacctcat caacgggctc tacgacgccc
18361 accgctccat ggccccgctc ctgcacctgg cctcccagat cccgtccagc gagatcgggc
18421 tcggcttctt ccaggagacc caccgacgac agctgttccg cgagtgcagt cactacagcg
18481 agctgatctc cagcccgaag cagatgcccc ggctgctcca gaccgccatc caacacgccc
18541 tcggccaggg cggcgtcagc gtcgctctgc tgcccggcga catcgcgga cagcccgcgc
18601 cgagggggc cgccgagacg gccctcgta cctccggg ccacgtccgg cccggcgacg
18661 aggagatcga ccggctggtc cggatgatcg acgacgccga caaggtcacc ctgtctcg
18721 gcagcggcac cgccggcgcg cagcgaggt tcattggagt cgccgggaag ctcaaggcac
18781 cggtcgggca cgccctgcgg ggcaaggagt tcattcagta cgacaaccg tacgacgtcg
18841 gcatgagcgg actgctcggc tacggcgccg cctacgaggc caccacgag tgcgacctgc
18901 tgctcctgat cggcaccgac ttcccgta acgcttct cccggacgac gtgaagatcg
18961 cccagatcga cgtgcggccc gagcacctgg gacggcgct caagctggac ctgcgggtgt
19021 gggcgacgc gcgggagacg ctgcgctgcc tgatccccg ggtcaaggag aagaagaacc
19081 gccgcttctt ggaccggatg ctgaagaagc acgcccagc cgatccacc cgagtacgtg gccgcctgc
19141 cgtacaccg caaggtcgac aagcacgtc cgatccacc cgagtacgtg gccgcctgc
19201 tggacgagat ggccgacgac gacgggtgt tcaccgtcga caccgggatg tgcaacgtct
19261 gggcggtcg ttacatctcg cccaacgggc cccgcccgat catcggttcc ttctccacg
19321 gttcgatggc gaacgcgctg ccatgcgga tcggcgccca gttcaccgac cggcgccgac
19381 aggtcgtgtc gatgtccggc gacggcggt tcacatgct gatggcgac ttctcacc
19441 tcgtccagca cgacctgccc gtgaagatcg tgctcttcaa caactcctc ctccgcatgg
19501 tcgagttgga gatgctcgt gcgggcctgc cctcgcacg ggtggccaac aagaaccgg
19561 acttcgccc cgtcgccgag gcctcgggc ccttcggcgt ccgggtggag aagcccagg
19621 acctggccg ggcgtgaag ggcggttca agcacaaggg tccgccttc gtcgacgtgg
19681 tcaccgaccc caacgcctg tccatcccgc cgaagatcag cgccgacatg gtcaccggct
19741 tcgcccgtgc cgctcgaag atcgctcgtg acggcggggt cggacgcatg ctccagatgg
19801 cccggtcgaa cctgcgcaac gtccccgc cgtgaccgta cggacagccg gttcgaaacc
19861 ggatgactgg cagggctcg aacgggcaag gctttcccc tgaagtgtt cgaccggagc
19921 gggacggacg gggggagacg acatcagcgc gccgacggg gacatgaagg agcaggcgg
19981 gccacagctc agggcgcggc tcggccgggc ggacctgcg gcggtgccc agggccggc
20041 ggcgtgcgt gagctgctgc ggcactggg cgggcccgg cagtcggagg tcgaggagt
20101 gctcgccagc gagctggtca ccaacgcgt cgtgcacacc gacgagggg cggctcctgac
20161 cgcgacggtc gggcgcgcg ccctcggggt cgaggtacgg gacttcgtgg gccggggccg
20221 gcagccccgg ccgcgcgcac cgcagcgga ggagagcacc aacggcagag ggtggtcct
20281 ggtcgagtgc ctgcgcgacg actggggcgt acagccctgc gaggtgggca agtcggtctg
20341 gttcgaactg ggcgcgggag cggaaaggga agcggaagc gcctgaggcg ccgggaagac
20401 cggcgcccca ggccatcggt atcgatacgg gggcgcgctc ggctcagccg aactgctgct
20461 ccaggctcct gagcttgccg tccaggaggt ccaggcgggg cagggcctgg gtgtcgctct
20521 ccgcggtgag gtcgacggtc cccacggaca tctgctcaga cccgctccgc acgggtgca
20581 gggaggggcc ccggtcgcg gcgggcaggt gctccggggt cgatatggca ggtccgcgg
20641 cggcgcggtt cgagcctcgc tcgacctgct gcacctcgac ctgacctga cggcgccgc
20701 gcccggggaa gcccggtgg ccgtggcccc ggtgatcgc cttgagctgc gcccgctcca
20761 gccgctcctg ctgcgcgcg cggagcttct tctcctcct ctggcgccg ctgcgcga
20821 cctcctcgac cgctcgtcc aggtcgca cgccctccag gagcatcagc gaccaggcgc
20881 ggtactgtc gcggggcgcg gcgagccagc ggacctgcg gatctgcggc agcggacgcg
20941 gcaccagacc ctgctcgcgc agggcgggcc ggcggtctg cttcagggcc cggtcgaaca
21001 gcaccgcgc cgacagggac atgcccgcga agaagtgcg ggcacccgc tgccccaggc
21061 cgcgcgcgcc gtgcacccag ttgaaccagg ccgcggcgaa ggcgaaggcc cacacgagta
21121 tccgcgagcc gagcgccgcg tcaccgtggc tggcctcgcg caccgcgagg acggagcaga

21181 acatcgccgc cccgtcgagc ccgaacggga ccaggtagtg ccagccgtcg ctccagccga
21241 ggttctgtct gccgaagccg accagaccgt ggaaggagag cgcccgccgc acccgccgac
21301 agcagaagag caggacgtag gagaagggtg cgtagagggc ctcttctgcg cggcgccgct
21361 cctccatgcg ctcccacgag tcgtccgcgc ccgcgtcctt ccccgaggaa cgcttgccgc
21421 gcgcgagcac cgccaccgcc gccagcatgc ccaggagcag cacggcgccc ggaagcagcc
21481 agttcagcga tatgtcggtc aatctcatct ggggtccctt gcagtgggat agggcgtaac
21541 gcccgcctatg gtggcccaaa cccagcggcc ctccaggggt ttccggggcaa gaggccgcca
21601 agggagtgcg gaggggggtg ccagggcggc gttctgtctg aactgccgcg tgagggcgcg
21661 gagttgagtt cgattaagac taccgtacg ggtggttcca cggaaagtcc ccgcgacgtg
21721 tgggggagtt gtggaacaac gcggacgatc ttagccgacg gcccgtcagc ccgcgcgggc
21781 gagcagcttg tccacgcggt cggcgctcga ggtccgcggg caggtggcgc aggtgtcctc
21841 ggggcgaggg gtgtagaaca tgcagcagct cgcccggtcc cgggtgtgca gcggggcgcc
21901 gtccggcccc ttccagttcac ggaaggccgc ctcccgacg tacggcctgg tggcgcccg
21961 cagcagcagc tccagctcgt gccgcgcccg ctctgtctcg ccgagcaggt gggcgacgta
22021 ccacaggccc tcgacgacct cgtcggtcgc catccccac agggcacgtc cgcgcgcgcg
22081 catccggggg ccgaatccgg ccagcaccgg ctccaggtgc tcggcgaccg cggcccgac
22141 ctccggcgcg agcgctcct cgtcggcgac caccggggcg ccgggcagcg cggcgcgggg
22201 gtcgccgggg aggcaggcga aggacgcggc acgcaccgcc atccggccga gggggagccc
22261 ggcggccggt cggctcgtag acacgtgggc cgccgggtgg cggggcaccg ggcggtgag
22321 gaaccacggc acggtgatca gcaggcaggg gggccacgcg taccggtgca ggccgaagct
22381 cgcgatcacg tccggccggg cccgtgccc gtagtccgc agcacctggg cctcgtccca
22441 ggcgaggaag gccgccagct ccgtcccgcc cgccggcagg gaggcgccg tgaccagcc
22501 gccgcgctc gggacctcgt cggcccgccc cagttcggtc acgcccagcc cggggaggac
22561 ctccgacagg cgcgtgtagg cgtccgtgag ggccgtacgg gcgacaggca tgcgggacca
22621 ccgatccacg tcggaacaac gggctgctta aagtaagcc ttaccttacc caatatgttc
22681 gggatttgaa ctgacggccc cgtcgcttat ggtgcttgac ggacaggtga caaccgcgcg
22741 taatgacccc aagtaccgac acgtcccaca ggaggacccg tgaagcacag cgcgcagggc
22801 gccgaagaa cgggtgtggg tacgggtacg gggcgggcgg ccgaagcggg gcgatcccg
22861 gtgcagccgg gagcggcgga cagggcacgg ggccgggccc ccggcggtgc ggagccggac
22921 ggccggcac gcggcgagca caccacggc gaaccgcccg tgccccgccc gcgggccctg
22981 gtccagcgtt cctccgtgcg cggacagatc ctccagcccc tgcgctccgc gctggtcacc
23041 ggtgaactgc ggcccggcga ggtctactcg gcaccggtgc tcggcgagcg gttcggcgct
23101 tcggccaccc ccgtgcggga ggccatgcag cagctcgccc tgagggcgcg cgtcaggtc
23161 gtccccaacc ggggtctccg cgtcctgaa cgcggtgacc gcgagctggc gcaactggcc
23221 gaggtccggg cctgatcga cctcccggtg tggctccggc tggcccgcac ggtgcccgcg
23281 gagcactggg cggagctgcg cccgtcgcgc gagggcaccg tccgggcccg gtcctccggc
23341 tgcccggcca cctacgccga ggccgaccgc gccttccacc gcgcgcgct cgcgctggcc
23401 ggcaacgaac agctcgtccg gatcgccggt gacgtgcacc gccgcgccc gtggccgccc
23461 gccggtcgc cctccgtgcg cggccgggccc gacctggtcg ccgacgcga ccagcacacc
23521 gcgctgctgg acgcgctgat cgcggcgac ttggacgtgg tcggggccct ggtcggggac
23581 cacttcacag tcacgtcgtg agccggaggc ccggcagggc gcccttcacc cgggcgcccg
23641 gacggggccc ccggaccgga cgtgccccg gttgggacgt gtcgggcccg ctcgtccgcc
23701 gggcttcgcc cggatcggcg ttcccgggccc ggggtcccgg gtcggcgctt ccgggcccgg
23761 gcctcggacg gggcgggcca cggaccgcgc gccccgggccc cgggcccctt tgagcgaagg
23821 ggcccggacc ttcttgcgcc cccctgctag cggacctcct gccgttccgt gcggtcgtg
23881 ggcacctgct tccccggctt ctccgacttc cccgtggcgc ggggtccggc gccctgcgg
23941 cgcccggggc ggggaaccgg gtccgcggtt tgcagctcga tggccagggg agaggcggtg
24001 aagaccgacg agtacgtgcc gacccccac ccgatcagca gcgcgagcgc gaagtccgtg
24061 agcgatcgt cggccaggat cgccagcgag gccaggatga gcaccgcgcc catgcccgtg
24121 ttgaccgtgc gggcgagggt ctgcaggatc gcgtcgttgg tgagccggtc gaaccggtc
24181 ttccgctcct tgcccagcag ttcccggatg cggtcgaaga gcaccaccga gtcgttccac
24241 gagtagccga tcacggtcag cagcgccgcc aggaacaccc cgtcgatcgg cttgcccagc
24301 caggcgaaca caccgaccag aatgaccacg tcgtgggcga gcgcgcgac cgcgcgctg
24361 ccgaacagca ggcggaaccg gaccgccagg tacgccagct gcgcggccag cgcagggccg
24421 agggcgatca gggcgctccc gcgcagctcc tcgccagggc tcgggcccga cagctcgtcg
24481 cggaccttct ccgtctcgcc gccagttcg gcgacggcct tcgtgacggt cgcgcctcgc
24541 gtgtcgggtc gctcctccgt gcgcaccgtg aggtcgcgct cgcgggagga ctggacgacg
24601 gcgcgcggga agcccgcgtc ggccaggggc tcccgcgccc ggtcgggatc gacctgggtg
24661 gcggtggagt actcgatgag ccgcccgcgg gtgaactcga tgccgaagtt caggccccgc
24721 accaggatgc cggagccccg caccacaagg acgatcagcg aggcgcgacg ccagcgccgc
24781 ggccggcgca tcaggacgg gtccgcggcg agcagggcgt caccgaccgg gccggtcgag
24841 gagatgccg tgatacggg gcggcggaag accgcggggc ggctcgcggc gaactcggcg
24901 agcaccggg tgatcaccag ggcgtgacc atcgaggcca gcacaccgat gccgaggggtg
24961 acgcccgaag ccttgaccgg gccggaggcc aggaagaaca gcagggcgcc ggcgatcagc
25021 gtggtgatgt tggagtcggc gatcgcgctg aaggcgctgc ggaaaccggc ggtcagcgcg

25081 gagcgcgtac tgggacgggt cggggccgcc tgttcctcgc gggcccgcgc gaagaccagg
25141 acgttggcgt ccaccgccat gccgatcgcc agcacgaagc ccgccaggcc cggcagggtc
25201 agggtcgcgc cgaccgcggc cagggcggcg taggagatca ggccgtagca gagcagggcc
25261 acggtcgcga gggcgcccat cagccggtag acgacgatga tgaacagcgc ggtcagcgcg
25321 gtgccgatga cggcgcccca ggcgcggcg tcatggcct cgtcgcccaag ggtggcgccg
25381 atggtgcgct gctcgatggt ctcgaccggc accggcaggg cggcgccctt gatgagcagg
25441 gccagttcgc gcgcctcggc gtcgtcgaag gagccggtga tctgcgtgga gccgcgggtg
25501 atgcccgcgc cgacaggacac ggaggggtcg acctgcggcg aggagatgat cttgtcgtcg
25561 agcacgatgg cgacccggcg cgtgggggtca ccggccgggt tgcaggccgc ctcgccggtc
25621 acctgcgccc agcggtcgct gccggagtc tgaagtcga cggtgacggt ccagccggcg
25681 ccgtttctgct ggtcgaagcg ggcgtcggcg ccttgacgt cctgaccggt cagcgtggcg
25741 gccttcagcc gcagggactg cccggactcg tccggcagca cccgctcgcc gtcctcggcg
25801 gtctcgggtc tgcctcgggt cgcgtcggcg ggcttctccg ccgcgcgag caccgagtgg
25861 acggtgagct gcgcgggtcg gccgagcacg tcggcgccct tcttcgggtc ctgtacgccc
25921 ggcagctcga cgacgacccg gttctcgccg gagcggacga tggtcggctc ggcgacgccg
25981 agggcgtcga tgcggcccg cagcacctcc acggtgcggt cggtgccctc ccgtcggcc
26041 ttggtggtct cggtcgactt ggtctccagg acgatctggg tgccgcccgc caggtcgagc
26101 cccaggcgga ccggcatggt caggcgatg gccacggccc cggccatcac ggccagcgcg
26161 acgaaggcgc ggacgttcag ggaacgggac cgcgtacggg aacgggagcg ggaaggggga
26221 cgggagcgtt tcaagacggg cctccggcg gcacgcccg gcggcgccg agcgcggccg
26281 cggcggaaga agatgcaggg aaggagtgt gtctcagggt ccggaggaga ccggtggcgc
26341 ccggccgctg tcgtgggtcg cgcggccggg ggagaccggc accgggccg gcgcgaacga
26401 cgtccgcgcg tggccgcagg gggcgccgc gcccccagg tggcggttg tcgccgggtg
26461 ctggtccggt gccgtcggac gcccgcccg ctggtcgtgc cgggctcag gctggcgccg
26521 gcagaccggc gtgcagccgt cgtcggcgtg gagtcgagg gccgggtgc gcacggccg cgggccaggc
26581 ggcgccggcc gccgcgagg gccgcgtcg accgacgtgc gcacggccg cgggccaggc
26641 cgggaccagg ggcagcagga cggtcagcag gacggtgagc acgctcgcga gccgtcgtcg
26701 catgcaccgt cccctttct ggtggatgag cgggcctggc ggggtgggga gcccggggct
26761 ccggccgggg agtgggcggc cagagccacg ggcggttcgg gggaccgtca gggctcgatg
26821 agcccggcgc ggatggcgta ccgggtcagc tccaggcggt cgcgcaggcc gagcttctgc
26881 agcaggttcg cccggtgccc ctggacggtc ttgatgctga tgaagaggat ctcgccgatc
26941 tccttcgagg agtggccctc ggcgacgagc ttgaggacct cctcctcgcg cggggtgagg
27001 atgtggtcgg aggtctctc gccgtgccgc acccggtcga ggtagttgcg gatgagcgcg
27061 gtcaccgcgc ccgggtacag gaacggctcg tcgcgcacgc cggcgccgca ggcggcgacg
27121 aggtcccggc cggcgaccga cttgagcacg tatccgcagg ccccgactt cagcgcctgg
27181 aacaggtaact gctcgttgtc gtgcacgtc agcatcagga tgcgcagccc cggtttcagc
27241 gcggccagtt cgcgggcccgc ctggaggccg gtgaggcgcg gcacgcgat gtccaggacg
27301 gcgaggtcgg cctcgtgggc ccgcgccatg tcgatggcct ccgcgccgtc cccggcctcg
27361 gccaccacct ccaggtccgg ctcccggtcg aggatgaggc gcacaccgcg gcgcaccagc
27421 gcgtggtcgt cggcgagcaa gatgcggatc ttcgaggggg cggacgccg ggggtgtgccg
27481 tacgcggacg tcgtggacgc ggacgtgggc gtcgtgggcg tcgtggacgc gccggcgtgg
27541 cggccgcgcg cggacggctc ggacggctcg ggcaggagac ggtcggacat ggtcgtgct
27601 acttctgtgg aagggtgcg gtgagccgga tgcgggtgcc ggcgcgggt gcccgagtga
27661 tgtcctgggc ggcgccgatg agcagggcc gttcccgcac gccgcggatg cggcgccct
27721 cgtgcgcgcg ctcgatggcg cgcgcgtcgt cggcgacggt gagcgtcacc gcggcgctcg
27781 cgtgggcccag gccacctcc aggcgtcgg cgtccgcgtg gcggcgccg ttggtcaggc
27841 tctcctgagc caccggtag aggaaccagt cggctcctcg gtccaggacc ggcagatcgg
27901 cgtcgaacac gcgcaccacc cgcagaccgg tgtgggtggc gaagtcgtgc gtgagcgagg
27961 acagcgcgct gatcagaccg aggtcgtcca ggacaccggg ccgcagccg cgcaccaggc
28021 ggcggaacct gtccaggctc tcccgggtga tctcctgcgc ctggtgcagc cgtcgcgca
28081 gcggctctct ggctcgtcc gccgcagcc ccaggaccag caggatcgcg gtcgtgctct
28141 gcccgaacct gtcgtgcagc tcttggcgca tgcggcgccg ttccgcctcc tgggccagca
28201 ggacgcgggc gctgctggtg gcccgctcgt tctccaggcg gtccagcatg gcgttgaagg
28261 tgcggatcag ctccgggacc tcgcgcgcgc cggagatcgg cagccgctgg ccggggcgca
28321 gcaggtcgac ggtggtcatc agcctggtga gccgctccag cggcgccagg cccagcgca
28381 gcagggcacc gttggccacc agcatgacgc ccatgccgac cacgaggacg acggcctcgg
28441 tcagcaggat cggcacggag acggtcaccg gagcccacag cagcagcgcc gtggcgaaagc
28501 ccagcaccac cgcgttgagc ccgaagatcc gccagaacag ggacaccggg ggtacgcctc
28561 cttcggcgac gtgttgattt ggtacgcctc tactgtcgggt catcgcgcac cctcgtggg
28621 agctgcggcc caggtagcga ccgagcaagc gaccagagtg ctctcgcgg accggtcgg
28681 tcgatggggc ccgaccccca tttccggccc gcggcatacc catatggccg ccccgcgca
28741 ccgcccagtg gatccgcac ccccgggcaa atgggtcctg cgcgcgatgg gtttcccgg
28801 gcggcccggc cagggttgag gacgaccggg tccgatcggc ccggcactcg gcataccgc
28861 agcaccttcc gagaaaagga cactcatcg tgtcgtcga cgcctccgc atcgaaaccc
28921 gcccgcagcg gctgaccgct cagcaggcgc gccagcgcc cgaacacgc cgcaacaccc

28981 gcgtcaccca gcttcaggcc ctcgccgaga gcggccaggc ggacgaccag ctgatgtccg
29041 cgcagaaggc cgcgatcgag cgggtgctca aggagatcga cgaggccttc gcccggtcg
29101 aggagggcac ctacggcgcc tgctcggtc gcggcaaggc cgtccccggg gagcgctgg
29161 agatcctccc ctacacgcgg tactgcgtcg cctgccagcg ccgcgcggcc gcctgacgcc
29221 ccgtcccacc gtcccaccgt ccccgctccag taccgtccct cccgccctgc ccaaggggtg
29281 aagtggtgaa caaccagatc atcggcgacc gcgacacgcg tctgccgccc ctgtcgcccg
29341 aggacctcgc cgcgctccgg gacaacctcc gggagcaacg gctgttccgc gaggagcaac

//

[Restrictions on Use](#) | [Write to the HelpDesk](#)
[NCBI](#) | [NLM](#) | [NIH](#)